

Головко В.А., Безобразова С.В.

НЕЙРОСЕТЕВОЙ ПОДХОД К ДЕТЕКТИРОВАНИЮ ЭПИЛЕПСИИ

ВВЕДЕНИЕ

Заболевание эпилепсией охватывает около 1% человечества [1,2] во всех возрастных группах. Эпилепсия – хроническое заболевание головного мозга человека, характеризующееся повторными припадками, которые возникают в результате чрезмерных нейронных разрядов (эпилептические припадки) и сопровождаются разнообразными клиническими и параклиническими симптомами [3]. Около 60 миллионов людей не имеют возможности жить нормальной, полноценной жизнью, так как могут принести вред себе и своим близким. Лечение эпилепсии на сегодняшний день производится преимущественно проведением хирургической операции по удалению очагов заболевания или назначением фармацевтических препаратов.

Для исследования и обнаружения эпилепсии широко используется электроэнцефалография, предметом изучения которой являются электрические явления в мозге человека. Электроэнцефалограмма (ЭЭГ) представляет собой запись суммарной электрической активности мозга, которая позволяет судить о его физиологической зрелости, функциональном состоянии, общемозговых расстройствах и их характере [3]. Таким образом, сигналы ЭЭГ весьма показательны для обнаружения изменений в работе мозга.

Для обнаружения эпилептических приступов по сигналам ЭЭГ в основном используются линейные (частотно-временные, математические и статистические) методы, которые не дают достаточно точных результатов [2].

Электрическая активность нейронов головного мозга характеризуется нелинейной хаотической динамикой [4-7]. В качестве меры хаоса может служить корреляционная размерность или старший показатель Ляпунова [5]. Корреляционная размерность характеризует количество степеней свободы динамической системы. Чем сложнее система, тем больше ее корреляционная размерность. Старший показатель Ляпунова характеризует среднюю скорость расхождения двух начальных состояний системы. Его положительное значение свидетельствует о хаотическом поведении. При эпилептических припадках наблюдается уменьшение корреляционной размерности и старшего показателя Ляпунова, что характеризует переход от хаоса к порядку. Физически это можно объяснить тем, что чрезмерные разряды эпилептических нейронов нейтрализуют активность других нейронов головного мозга. В результате получается более упорядоченная картина электрической активности нейронов головного мозга. Для детектирования эпилепсии необходимо разработать робастные методы определения хаотичности ЭЭГ сигналов.

В данной статье исследуется нейросетевой подход расчета старшего показателя Ляпунова для определения эпилепсии. В разделе 1 статьи описывается метод независимых компонент для предварительной обработки ЭЭГ данных. Раздел 2 посвящен нейросетевому подходу для определения Старшего показателя Ляпунова. В разделе 3 приводятся результаты экспериментальных исследований по детектированию эпилепсии на основе ЭЭГ данных.

1. ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ИСХОДНЫХ СИГНАЛОВ

Как уже отмечалось, ЭЭГ характеризует электрическую активность головного мозга. При этом электрическая активность, снятая с одного скальпового электрода, представляет собой суммарную активность группы нейронов. Следует так-

же отметить, что энцефалограмма отражает не только сигналы коры головного мозга, но и другие сигналы, которые называются артефактами. К последним относятся различные погрешности аппаратуры и физиологические артефакты: электрическая активность сердца, движение глаз, мышечные сокращения и т.д.

Предварительная обработка ЭЭГ заключается в выделении сигналов, несущих полезную информацию об исследуемом процессе. Это достигается путем сокращения артефактов и шумов, а также выделения исходных источников сигнала.

Эффективным средством предварительной обработки сигналов является метод независимых компонент (ICA) [8]. Он базируется на предположении, что ЭЭГ данные, регистрируемые электродами, представляют собой линейную смесь независимых сигналов от различных источников. Пусть имеется n источников исходных сигналов $S_i(t)$ ($i=1,2,\dots,n$), которые представляют собой электрическую активность нейронов головного мозга и физиологические артефакты. В результате линейной комбинации статистически независимых сигналов ЭЭГ электроды будут фиксировать совокупность смешанных сигналов (рис. 1).

$$X_j(t) = \sum_{i=1}^n S_i(t)w_{ij}, \quad (1)$$

где $[w_{ij}]$, $i = \overline{1, n}$, $j = \overline{1, m}$ - неизвестная матрица смеси.

Задача состоит в том, чтобы на основе смешанных сигналов $X_j(t)$, $j = \overline{1, m}$ найти исходные источники сигналов:

$$S_i(t) = \sum_{j=1}^m v_{ji}X_j(t). \quad (2)$$

Здесь предполагается, что количество исходных источников сигналов n должно быть не меньше, чем количество смешанных сигналов m . Решение данной задачи заключается в выделении из смеси сигналов статически независимых компонент. Для этого можно использовать в качестве целевой функции различные критерии, например, максимизация негэнтропии или момента четвертого порядка (эксцесс) [8]. Как правило, перед применением метода независимых компонент производится отбеливание смешанных данных путем использования метода главных компонент.

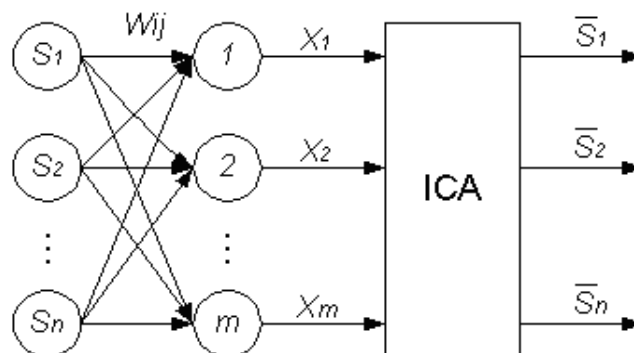


Рис. 1. Диаграмма смешивания и разделения сигналов.

В результате применения такого способа предварительной

Безобразова С.В., аспирант каф. интеллектуальных информационных технологий Брестского государственного технического университета.

Беларусь, БрГТУ, 224017, Беларусь, г. Брест, ул. Московская, 267.

обработки данных можно избавиться от шумов и артефактов.

2. РАСЧЕТ СТАРШЕГО ПОКАЗАТЕЛЯ ЛЯПУНОВА

Старший показатель Ляпунова характеризует степень экспоненциального расхождения близких траекторий. Наличие у системы положительной экспоненты Ляпунова свидетельствует о том, что любые две близкие траектории быстро расходятся с течением времени, то есть имеет место чувствительность к значениям начальных условий.

Рассмотрим определение старшего показателя Ляпунова на базе нейронной сети [9-11].

Ключевой идеей данного метода является вычисление при помощи прогнозирующей нейронной сети расхождения двух близлежащих траекторий на n шагов вперед.

Нейронная сеть будет состоять из $k \geq m - 1$ входных нейронов, p скрытых и одного выходного нейронного элемента (рис. 2). Здесь m – размерность пространства вложения.

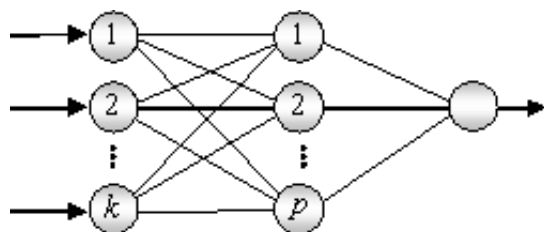


Рис. 2. Прогнозирующая нейронная сеть.

Вначале необходимо обучить такую нейронную сеть прогнозированию в соответствии с методом скользящего окна:

$$x(t+i\tau) = F(x(t+(i-1)\tau), x(t+(i-2)\tau), \dots, x(t+(i-k)\tau)), \quad i = \overline{1, n}, \quad (3)$$

где τ – временная задержка.

После обучения сети легко осуществить эволюцию двух точек на фазовой траектории, используя итерационный подход.

Таким образом, ключевой идеей предлагаемого метода является вычисление при помощи прогнозирующей нейронной сети расхождения двух близлежащих траекторий на n шагов вперед, используя итерационный подход. Эта процедура может быть представлена следующим алгоритмом:

1. Обучаем нейронную сеть на прогнозирование по методу скользящего окна.
2. Выбираем любую точку $x(t)$ из обучающей выборки и формируем следующий набор данных: $\{x(t), x(t-\tau), \dots, x(t-(k-1)\tau)\}$, где k – размер окна.
3. Вычисляем $\{x(t+\tau), x(t+2\tau), \dots, x(t+n\tau)\}$ используя многошаговый прогноз $x(t+i\tau) = F(x(t+(i-1)\tau), x(t+(i-2)\tau), \dots, x(t+(i-k)\tau))$, где $i = \overline{1, n}$, F – нелинейная функция.
4. Вычисляем $x'(t) = x(t) + d_0$, где $d_0 \approx 10^{-8}$ и, подавая на сеть $\{x'(t), x(t-\tau), \dots, x(t-(k-1)\tau)\}$, повторяем шаг 3 для получения $x'(t+i\tau)$, $i = \overline{1, n}$.
5. Оцениваем $\ln d_i = \ln |x'(t+i\tau) - x(t+i\tau)|$, $i = \overline{1, n}$ и выбираем только точки, где $\ln d < 0$.
6. Строим график $\ln(d_n)$ от n .
7. Строим прямую регрессии для выбранных точек и вычисляем её наклон, который равен наибольшему показателю Ляпунова.

Данный метод позволяет достаточно просто вычислить старший показатель Ляпунова при малом объеме экспериментальных данных.

3. ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ

Для исследования мы использовали реальные данные, взятые на сайте [12]. Эти данные представляют собой чистые (отфильтрованные от помех и шумов) сигналы ЭЭГ с различным характером и насыщенностью эпилептических событий. Сигнал, изображенный на рисунке 3 содержит эпилептическое событие в интервале (1.15 – 1.35) секунды.

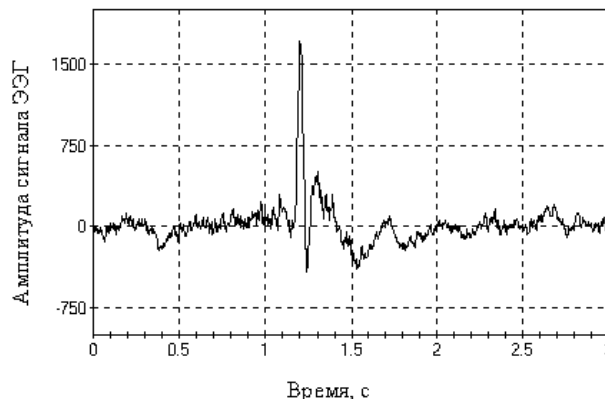


Рис. 3. Сигнал электроэнцефалограммы, содержащий одно эпилептическое событие.

Такие сигналы ЭЭГ позволяют наглядно определить эпилепсию и проверить результаты, получаемые при их анализе. Исходные данные, сигналы ЭЭГ, описаны в виде текстовых файлов, содержащих отсчеты амплитуды сигнала через каждые 0,03 секунды.

Критерием обнаружения эпилепсии выступает старший показатель Ляпунова:

$$\begin{cases} \lambda > 0, \text{ нормальное состояние;} \\ \lambda \leq 0, \text{ эпилепсия.} \end{cases} \quad (4)$$

Для анализа данных будем использовать расчет старшего показателя Ляпунова при помощи прогнозирующей нейронной сети. Предлагается использовать многослойный перцептрон (MLP) с одним обрабатывающим слоем, сигмоидной функцией активации в скрытом слое и линейной – в выходном. Обучение нейронной сети производится на основе данных ЭЭГ в соответствии с алгоритмом обратного распространения ошибки. Перед расчетом показателя определяются размер пространства вложения и временная задержка исследуемых данных (для сигнала, изображенного на рис. 3: $\tau = 4$, $m = 7$).

Расчет показателя λ осуществляется в каждой точке обучающей выборки. В результате получается зависимость $\lambda(t)$, которая представляет собой детерминированный ряд:

$$\lambda(t) = (\lambda_1, \lambda_2, \dots, \lambda_L), \quad (5)$$

где L – размер обучающей выборки.

Используя такую зависимость, сложно производить идентификацию эпилепсии, так как она очень нестабильна. Для уточнения результатов производится усреднение $\lambda(t)$ на n точек ($n \approx 10-20$), при этом устраняется излишняя скачкообразность результатов, что позволяет избежать ложной идентификации эпилепсии:

$$\bar{\lambda}_k = \frac{\lambda_k + \lambda_{k+1} + \dots + \lambda_{k+n-1}}{n}. \quad (6)$$

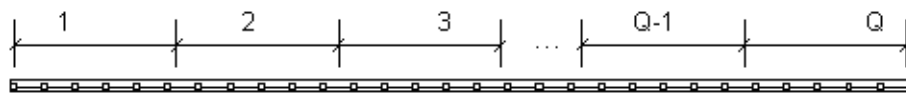


Рис. 4. Деление исходных данных на обучающие выборки последовательно, Q – общее количество обучающих выборок.

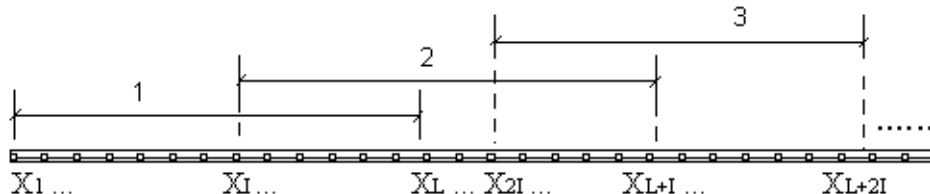


Рис. 5. Задание отрезков обучения L и прогнозирования I.

По полученной усредненной зависимости $\bar{\lambda}(t)$ можно судить о наличии эпилептических событий в сигналах, пользуясь критерием (4).

Большая погрешность детектирования аномалий вызвана тем, что обучение НС производится на L точках, а прогноз осуществляется на $L+H$ точек. Таким образом, при прогнозировании расчет показателя λ значительно снижает точность на последних $2H$ точках. Вероятно, это происходит из-за того, что выборки при обучении НС выбираются последовательно (рис. 4).

Для устранения этого явления можно предложить следующие действия:

- Разбиение исходных данных на обучающие выборки будет производиться с наложением друг на друга.
- Прогнозирование будет осуществляться не на всей выборке (L точек), а на первые I точек (рис. 5). Следующая выборка будет формироваться с $I+1$ точки.

Недостатком данного метода является увеличение длительности работы программы за счет увеличения количества обучающих выборок. Однако такой способ позволяет избежать неправильной идентификации эпилепсии, так как участки с большой погрешностью не будут участвовать в прогнозировании ЭЭГ данных.

При анализе сигнала, который показан на рисунке 6, получены следующие результаты:

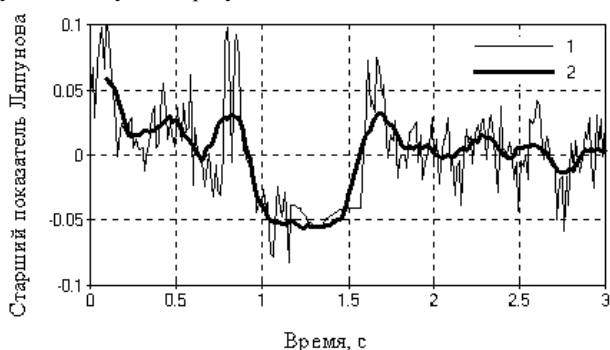


Рис. 6. Расчет старшего показателя Ляпунова λ на всей обучающей выборке: 1 - зависимость $\lambda(t)$, 2 - усредненная зависимость.

- а. Расчет старшего показателя Ляпунова производится на протяжении всей обучающей выборки (30 точек). Если рассматривать получаемую зависимость $\lambda(t)$ (см. рисунок 6), то, очевидно, что она носит нестабильный характер, содержит много «скачков», которые могут стать ложно-

определенными эпилептическими событиями. График усредненной зависимости позволяет определить не только присутствие эпилепсии, но и приблизительный участок наступления эпилептического приступа (0,9-1,6) секунды. Однако в данном случае получается большая погрешность интервала эпилепсии (приблизительно 0,25 секунды).

- б. Расчет старшего показателя Ляпунова производится на первые M точек обучающей выборки ($M = 10$). Будем рассматривать только усредненную зависимость. Из рисунка 7 видно, что точность детектирования эпилепсии значительно увеличилась, так как отсутствуют ложные переходы λ в отрицательные значения. В результате, определяется одно эпилептическое событие в интервале (1,1-1,3) секунды. Погрешность составила 0,05 секунды.

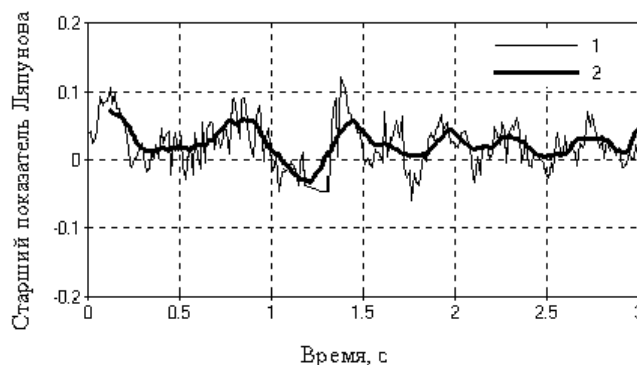


Рис. 7. Расчет старшего показателя Ляпунова λ на первые M точек обучающей выборки: 1- зависимость $\lambda(t)$, 2- усредненная зависимость.

Исследования также проводились для сигнала, содержащего девять эпилептических событий. Получены аналогичные результаты, которые приведены в таблице 1 и подтверждают сделанные предположения.

ВЫВОДЫ

1. Разработан подход для обнаружения эпилептических приступов на основе расчета старшего показателя Ляпунова. В качестве исходных данных используются сигналы ЭЭГ. Данный метод позволяет определять как одиночные приступы, так и целый ряд следующих один за другим эпилептических событий.
2. Предложена методика по устранению ложно определяемых эпилептических событий, которая позволяет с высокой точностью определять не только момент наступления эпилепсии, но и длительность приступа. Для данных, взятых через 0,003 секунды, погрешность составила 0.05 секунды.

Таблица 1. Результаты детекции эпилептических событий

№ п/п	Эпилептические события (интервал t , с)	Расчет λ на всей обучающей выборке (интервал t' , с)	Расчет λ на первые $M = 10$ точек (интервал t'' , с)	Погрешность $\Delta t'$, с	Погрешность $\Delta t''$, с
1.	0,20 – 0,30	0,25 – 0,45	0,20 – 0,35	0,15	0,05
2.	0,50 – 0,60	0,45 – 0,55	0,50 – 0,60	0,05	0
3.	0,80 – 0,90	0,70 – 0,85	0,85 – 0,90	0,1	0,05
4.	1,10 – 1,20	0,90 – 1,20	1,05 – 1,20	0,2	0,05
5.	1,35 – 1,45	1,30 – 1,55	1,30 – 1,50	0,1	0,05
6.	1,60 – 1,70	1,65 – 1,85	1,65 – 1,75	0,15	0,05
7.	1,95 – 2,05	1,90 – 2,10	1,95 – 2,05	0,05	0
8.	-	2,15 – 2,25	-	-	-
9.	2,20 – 2,30	2,30 – 2,45	2,25 – 2,35	0,15	0,05
10.	2,55 – 2,65	2,55 – 2,70	2,55 – 2,65	0,05	0
Итоговая максимальная погрешность:				0,2	0,05

3. Исследовано применение старшего показателя Ляпунова в качестве критерия детекции аномалий, возникающих в сигналах ЭЭГ. Определение наличия эпилепсии в сигнале производится согласно условию (4).

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- Verdes P.F, Deco G, Obradovic D, Dubé L.J, Hopfengaertner R, Stefan H. Detection and prediction of epileptic seizures: a patient's case study.- www.tecn.upf.es/~gdeco/pubeng.html, 2000
- Litt B, Echaz J. Prediction of epileptic seizures: review. - <http://www.ncbi.nlm.nih.gov/entrez/>, 2002
- Карлов В.А Эпилепсия. – М.: Медицина, 1990. - 336с.
- Babloyantz A., Destexhe A. Low dimensional chaos in an instance of epilepsy. Proc. Nat. Acad. Sci. USA, 1986.
- Toward a Neurodynamical Understanding of Ictogenesis // Epilepsia, v.uu, №1, 2003, - P. 30-43.
- Moser H., Weber B., Wieser H. Electroencephalograms in epilepsy: analysis and seizure prediction within the framework of Lyapunov theory. Physiol. D, 1999, - P. 130, 291-305.
- Sackellares J.Ch, Iasemidis L.D, Shiau D. Epilepsy when chaos fail. Singapore: Word Scientific, 1990.
- Huvaerinen A., Oja E. Independent component analysis: algorithms and applications // Neural Networks, №13, 2000, - P. 411-430.
- Головки В.А., Чумерин Н.Ю. Нейросетевые методы определения спектра Ляпунова хаотических процессов // Нейрокомпьютеры: разработка и применение, 2004. – №1.
- Vladimir Golovko, Nikolay Maniakov, Alexander Doudkin. Application of Neural Networks Techniques to Chaotic Signal Processing//Optical Memory and Neural Networks, vol.13, Number 4, 2004, - P. 195-215.
- Головки В.А. Нейросетевые методы обработки хаотических процессов// В книге «Лекции по Нейроинформатике». – М.: МИФИ, 2005. – С. 43-88.
- Данные электроэнцефалограмм - <http://republica.pl>, 2002.

УДК 004.8

Кочурко П.А.

СОВОКУПНОСТЬ ДЕТЕКТОРОВ НА ОСНОВЕ РЕЦИРКУЛЯЦИОННЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РАСПОЗНАВАНИЯ КЛАССА СЕТЕВЫХ АТАК

1. ВВЕДЕНИЕ

Непрекращающееся распространение применения информационных технологий во всех сферах человеческой деятельности постоянно ставит новые требования к уровню защищённости информационных систем. Стандартным компонентом инфраструктуры сетевой безопасности уже стали системы обнаружения атак (СОА) – аппаратно-программные комплексы, реализующие совокупность мер по обнаружению, определению и распознаванию вредоносных сетевых воздействий и реакции на эти воздействия – прерывание сетевой активности, оповещение персонала, ответственного за информационную безопасность. Несмотря на то, что существуют и постоянно появляются новые методы анализа сетевой активности с помощью различных технологий добычи данных («data mining») [1], основной технологией обнаружения атак по-прежнему является сигнатурный поиск. Его основной недостаток – недостаточная гибкость при обнаружении модифицированных атак [2]. Значительно лучшие результаты при определении модифицированных и новых атак способны показать системы, использующие искусственные нейронные сети [3-11]. Искусственные нейронные сети (ИНС) имеют потенциал для решения большого количества проблем, охва-

тываемых другими современными подходами к обнаружению атак. ИНС были заявлены в качестве альтернативы компонентам статистического анализа систем выявления аномалий. Нейросети были специально предложены для того, чтобы идентифицировать типичные характеристики пользователей системы и статистически значимые отклонения от установленного режима работы пользователя [2].

В данной работе рассматривается метод распознавания класса атак на основе анализа сетевого трафика. Обучение и тестирование ИНС производилось на выборке KDD'99, содержащей записи о TCP-соединениях, включающих 41 параметр, полученные из обработанной базы данных DARPA 1998 Intrusion detection evaluation [12]. Данная выборка включает нормальные соединения, а также атаки 23 типов, принадлежащие к четырём классам: DOS – «denial-of-service» - отказ в обслуживании, например, Syn-лавина; U2R – неавторизованное получение привилегий goot на данной системе, например, различные атаки «переполнения буфера»; R2L – неавторизованный доступ с удалённой системы, например, подбор пароля; Probe – наблюдение и другое зондирование, разведка, например, сканирование портов.

Кочурко Павел Анатольевич, аспирант каф. интеллектуальных информационных технологий Брестского государственного технического университета.

Беларусь, БрГТУ, 224017, Беларусь, г. Брест, ул. Московская, 267.

Физика, математика, информатика